Appendix 24. OCSE Predictive Model Example

This short paper describes using database information and statistical tools that will provide managers with predictive insight into child support collections. The paper provides an econometric basis for collection behavior and a range of statistical techniques to accurately predict collection success. Using such tools can provide decision-support services to caseworkers and managers at all levels of Child Support Enforcement Management.

Child support (CS) payments may be considered a routine monthly expense. If this underlying premise is correct, the child-support payment would behave the same way as the price of a product in the marketplace. This market behavior can provide predictability in the collection of CS payments in much the same manner we can predict that consumers will purchase a product based on varying prices for the product. It is important to consider this "market behavior predictability" in collection management of CS payments in order to minimize payments that are in arrears—or overdue.

The fundamental market behavior that must be considered is called *elasticity*. To understand elasticity, and its inverse, inelasticity, you need only consider your own experience in the marketplace.

Inelasticity means that, as a consumer, you will tolerate significant price changes in items you do not spend much on. If, for example, a pack of gum costs 10 cents and the price is increased to 15 cents, you may note the increase of 5 cents, a 50% increase. However, you probably won't reduce the number of packs you purchase by 50%. Chewing gum, in this price range, is inelastic. For inelastic products, large price changes do not affect the quantity of product sold.

Elasticity, however, means consumers may not tolerate steep price increases. Here are two examples: The first involves luxury yachts. One year, a 10% price increase occurred on expensive models as the result of a luxury tax, so a $200,000 yacht suddenly cost $220,000. Yacht sales plummeted, shipbuilders went bankrupt, and overall tax collections declined. Consumers withdrew from the market because they declined to accept the price increase. Yachts are elastic.

The second example, involving gasoline, illustrates that elasticity is not instantaneous (a conclusion we might reach if we only considered yachts). The price of gasoline may increase 100%, but near-term demand may not decline as it did in the yacht example, because leaving the market takes time. To leave the market, consumers will have to find other modes of transportation, buy more fuel-efficient cars, develop alternative fuels for cars, move closer to their work, or change their lifestyles. Consumers have few options and cannot leave the marketplace immediately but can over time. Therefore, elasticity of demand may have a time component.

CS payments, like a consumer product, may behave according to the elasticity of demand. The larger the percentage of a consumer's income that is spent on CS payments, the more elastic the demand. This means that a small percentage change in payment (either up or down) will result in a large increase or decrease in arearrs, or overdue payments. Increasing the burden by 5% will lead to greater than 5% in burden accepted.

The following explanation applies econometric theory to CS payment collection:

CS payments are considered the same as purchasing a product. The larger a monthly payment as a percentage of monthly income, the more likely payment will be considered to cost a lot of money. Therefore, the consumer will react negatively because demand will be elastic. The willingness of the consumer to pay for the product (child support) will decline significantly as CS payments increase as a percentage of income.

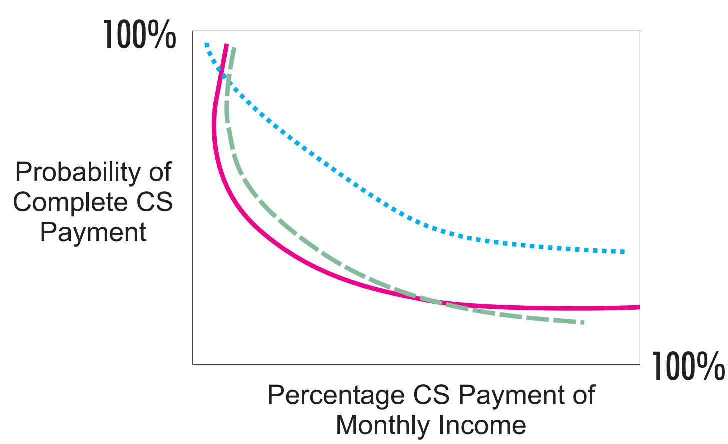

Figure 24-1. Income Effect on CS Payment

Figure 24-1 shows three curves that may indicate the behavior of those making CS payments in relation to how large a percentage of their monthly income must be devoted to CS costs. Using collected data, caseworkers and managers can determine what the curve is for their community, region, state, and income level. With this data, decision programs can be established to determine collection strategies, impact of proposed changes on CS payments, efficiency of collections by area and income, and other issues determined by CSE management. Existing data can be used to determine elasticity and the probability of collecting any or all of the CS payment.

The following is a suggested simple methodology that can be used to build data for management decisions.

1. Data has to be linked with: 1. Monthly income, 2. Amount of CS payment required, 3. Amount of CS payment paid, 4. ZIP code of payee.
2. Determine and maintain the percentage of CS payment to monthly income for each data set: This is *amount of CS payment required/amount of monthly income*. Also determine and maintain percentage of CS collected, which is *CS paid (collected)/CS payment required* for each data set.
3. Match the data with the items in step 1 so that for each set of data, the elements include: 1. Monthly income, 2. Amount of CS payment required, 3. Amount of CS payment provided, 4. ZIP code of payee, 5. Percentage of CS payment to income, and 6. Percentage of CS payment collected.
4. For each ZIP code, or bundle of ZIP codes, take the data and order it in sequence. One set of data may be by *percentage of child-support payment to income* (from smallest percentage to largest percentage). Another order statistic may be *amount of CS collected/required child-support payments* or *percentage of CS payment collected*. Ordering the data will enable the statistics to be more meaningful.

When using data from a sample set (e.g., database) to make decisions, it is essential that outlying/extreme data points do not detract from the actual majority of the population. Two approaches can be taken to establish population characteristics: the median and the mean. Half the population will have values less than the median, and half will have values that exceed it.

So in continuing with our example, let the amount of *CS collected/required CS payments* or *percentage of CS payment collected* be placed in order. In this example, there are 11 data points (0,0,30,50,75,77,80,82,83,90, 100). Once the data is placed in order, it is easy to see that 5 values are below the median and 5 values are above the median, which is 77. In small samples, the median can quickly be determined. In large samples, it is *(n+1)/2* for odd lots and for even lots it is the average of the order values of the *n/2* and (*n+2)/2* data points.

The mean is the average of the data elements. In our example above, it is the sum of the values divided by the number of points (in this case 11). The mean is 60.33. In the sample, the value is highly skewed by the two data elements that are 0. The median is a better descriptive statistic for data sets that are potentially skewed.

The two statistics—median and mean—describe the center of the data. What statistic best describes where the majority of the population lies about the center? This is sometimes referred to as the range or deviation about the center. This concept is illustrated in **figure 24-2**.

Percent Payment of CS	Data Set A	Data Set B	Data Set C
0-25%	4	2	2
25-50%	4	8	6
50-75%	4	4	6
75-100%	4	2	2
Range	0-100	0-100	0-100

Figure 24-2. Median vs. Mean

Looking at this set of data, the median and mean cannot be determined, but if there is a value in each set at the extreme values (0 and 100), the range is the same for each data set. The deviation has to be calculated from the individual values of each data set. The deviation or spread around the mean or median will provide an indicator of the shape of the data distribution. In data set A, it appears to be uniformly spread; in data set B, it is bunched about the 25-50% grouping; and in data set C, it centers around the 25-75% grouping. Conceivably, all three data sets could have the same mean and median.

When the median is used as the descriptive statistic for the center, we can determine with confidence that the median for the population based on the sample lies between two of the data points of our sample. Using the previous example of ordered values (0,0,30,50,75,77,80,82,83,90,100) where the median was 77, it can be said with 95% confidence that the median will be located between two values. This is $(n+1)/2 \pm \sqrt{n}$ located by using ordered statistics. Since n=11, the formula is 6 + square root of 11, which is 2 and 10. The second ordered statistic is 0, and the tenth is 90. Therefore, 95% confidence that the median is between those two values. As *n* increases, that interval (0-90) will shrink and the confidence level (95%) will remain the same.

Using the data collected and using descriptive ordered statistics (median) and the concept of price elasticity for CS payment, a DSS can be constructed for CSE management. Grouping data by geographical area (ZIP code) and adjacent ZIP codes increases fidelity of population estimates. In **figure 24-3**, the data is by adjacent ZIP codes.

Rec.#	ZIP Code	Monthly Income-$	CS Payment	% CS Payment/Income	CS Collected	% CS Coll./Payment	% CS Coll./Inc.
1	22181	3500	1200	34.3	1100	91.6	31.4
2	22181	1900	900	47.3	440	48.8	23.7
3	22180	1750	850	48.5	475	55.8	28.1
4	22180	1675	850	50.7	420	48	25.1
5	22180	1850	750	40.5	570	76	30
6	22182	2100	600	28.7	600	100	28.7
7	22182	1950	635	32.5	590	92	30
8	22182	1780	450	25.2	450	100	25.2
9	22179	1200	450	37.5	300	66	25
10	22179	1350	500	37.0	325	65	24.7
11	22179	1400	575	41.	250	43.4	17.8

Figure 24-3. Grouping Data by Geographical Area

Rather than focusing on the *percentage CS collected/payment* data, which indicates a range from 100% to 43.4% (56.6%), use the median of 66. That indicates that 66% is the median collection rate. Now look at the data of percentage collected/income. This indicates the percentage of income that is devoted to CS payments. That median is 25.2% with a range from 31.4% to 17.8% (13.6%). This indicates that there is price tolerance for CS payments for this range of incomes ($1,200-$3,500) per month, with a median of around 25.2% of the monthly income for CS payments. Thus, if you were managing the cases shown in **figure 24-3**, there are only two cases significantly below the median (Rec. # 2-23.7% and #11 17.8%) that could be potentially raised to increase CS payments to approach the median, an increase of $35 and $100 dollars. Unless monthly income increases in other cases, the potential for increased CS collections is small.

This data can be used to focus on cases, determine whether increases in CS payments can realistically be supported (collected) if no commensurate increase in monthly income occurs, track monthly income to determine if collection percentages increase can occur, and make other management decisions.

Keeping data current is important to detect changes over time. If monthly incomes rise, some percentage should go to CS payments—especially as the CS payment percentages continue to be below elasticity thresholds.